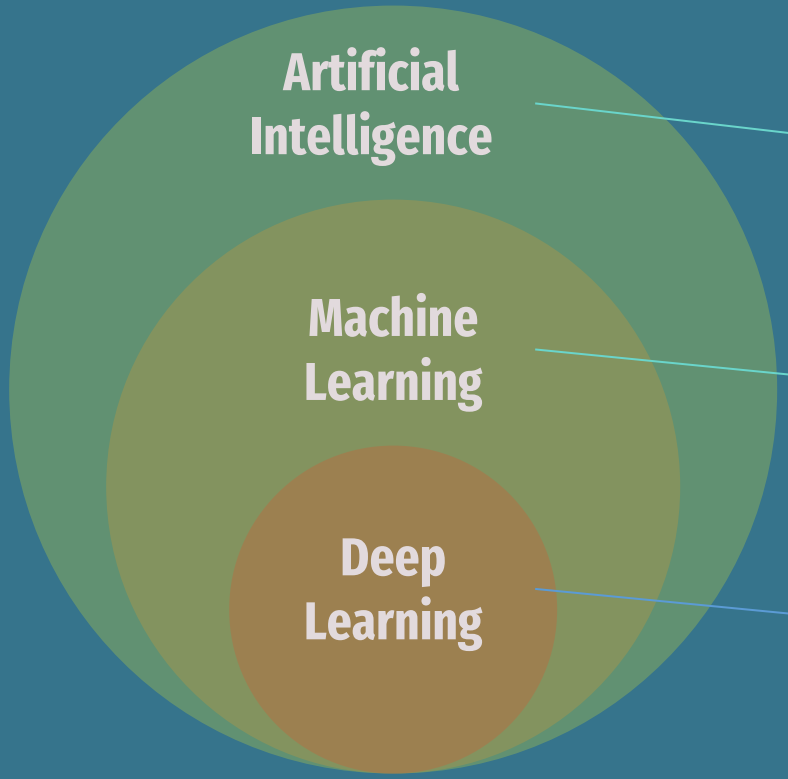
Introduction to Machine Learning

Bassem Ben Hamed
bassem.benhamed@enetcom.usf.tn

DIGITAL INNOVATION
PARTNER

AI vs ML vs DL



The theory and development of computer systems able to perform tasks normally requiring human intelligence

A field of study that gives computers the ability to learn without being explicitly programmed

A type of machine learning based on artificial neural networks

Use cases for ML

Healthcare

- Medical Imagining and diagnostics
- Personalized medicine
- Predictive approach to treatment

Finance

- Algorithmic trading
- Fraud detection and prevention
- Portfolio management

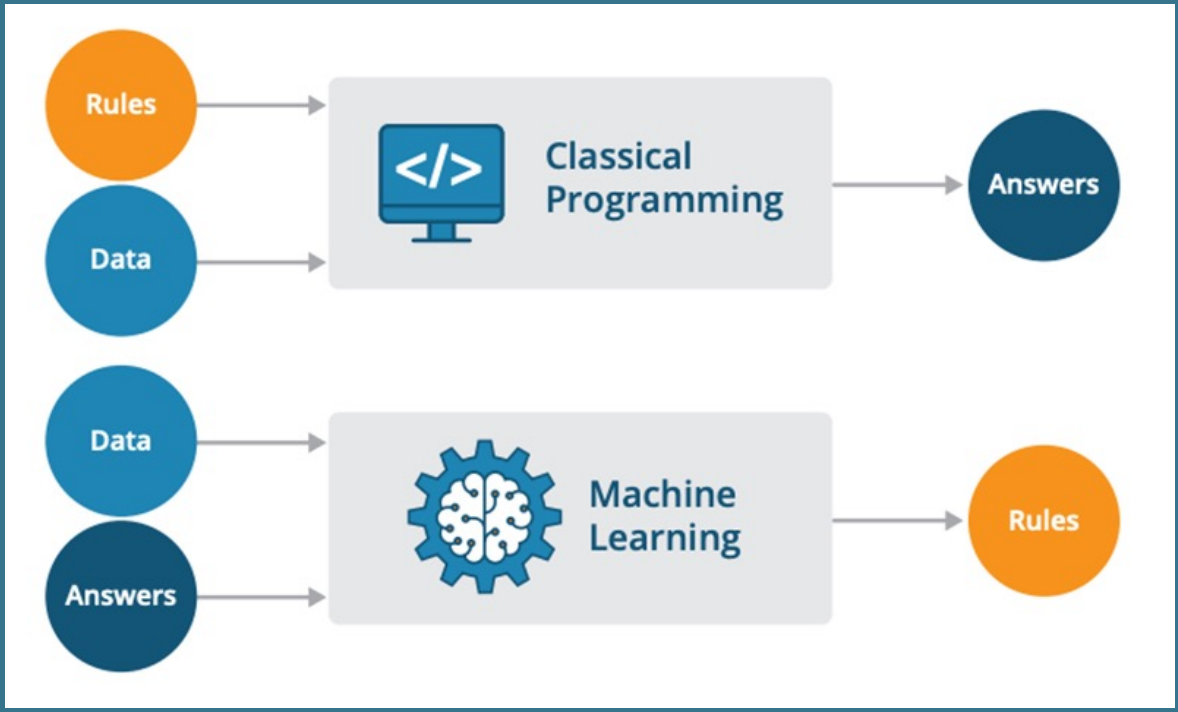
Telecom

- Anomaly detection
- Predictive maintenance
- Churn prediction

Manufacturing

- Energy consumption forecasting
- Predictive quality and yield
- Supply chain management

Traditional Programming vs ML



Machine Learning: Field of study gives computers the ability to learn without being explicitly programmed.

Arthur Samuel

What is ML?

Learning is any process by which a system improves performance from experience.

Herbet Simon

Machine Learning is the study of algorithms that

- Imporve their performance *P*
- At some task *T*
- With experience *E*

As well-defined learning task is given by <*P*,*T*,*E*>.

Tom Mitchell

Task vs Performance vs Experience

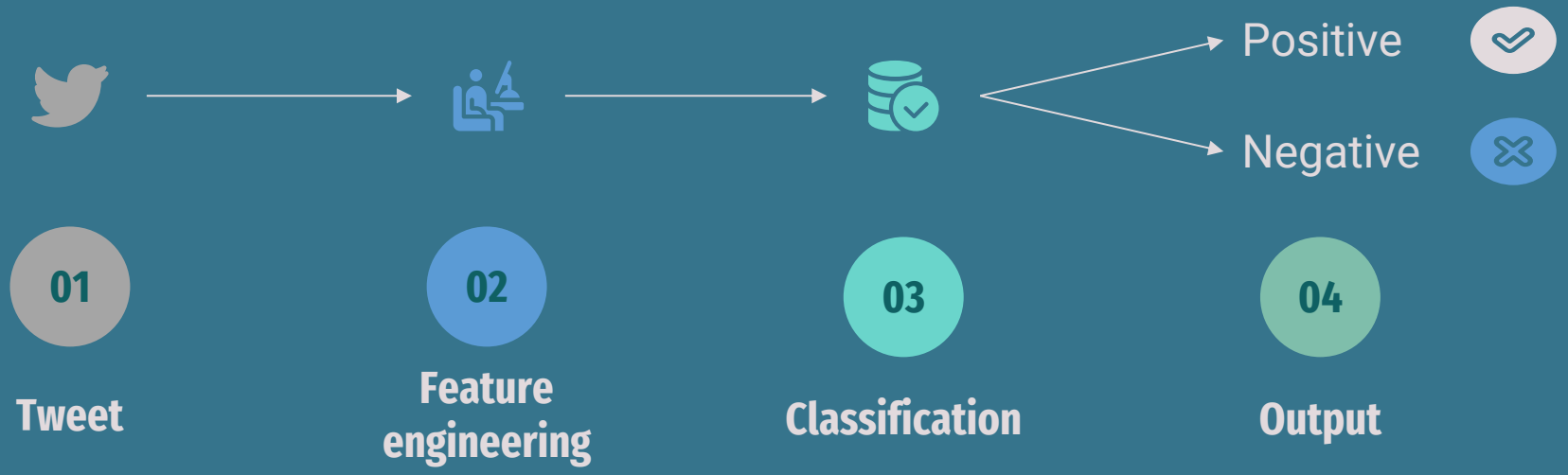
Improve on task T, with respect to performance metric P, based on experience E

T: Recognizing hand-written words
P: Percentage of words corrected classified
E: Database of human-labeled images of handwritten words

T: Categorize email messages as spam or legitime
P: Percentage of email messages correctly classified
E: Database of emails, some with human-given labels

T: Playing checkers
P: Percentage of games won against an arbitrary opponent
E: Playing practice games against itself

Twitter sentiment analysis example



Types of Learning

Supervised (inductive) learning

Given training data + desired outputs (labels)

Unsupervised learning

Given training data (without desired outputs)

Semi-supervised learning

Given training data + a few of desired outputs

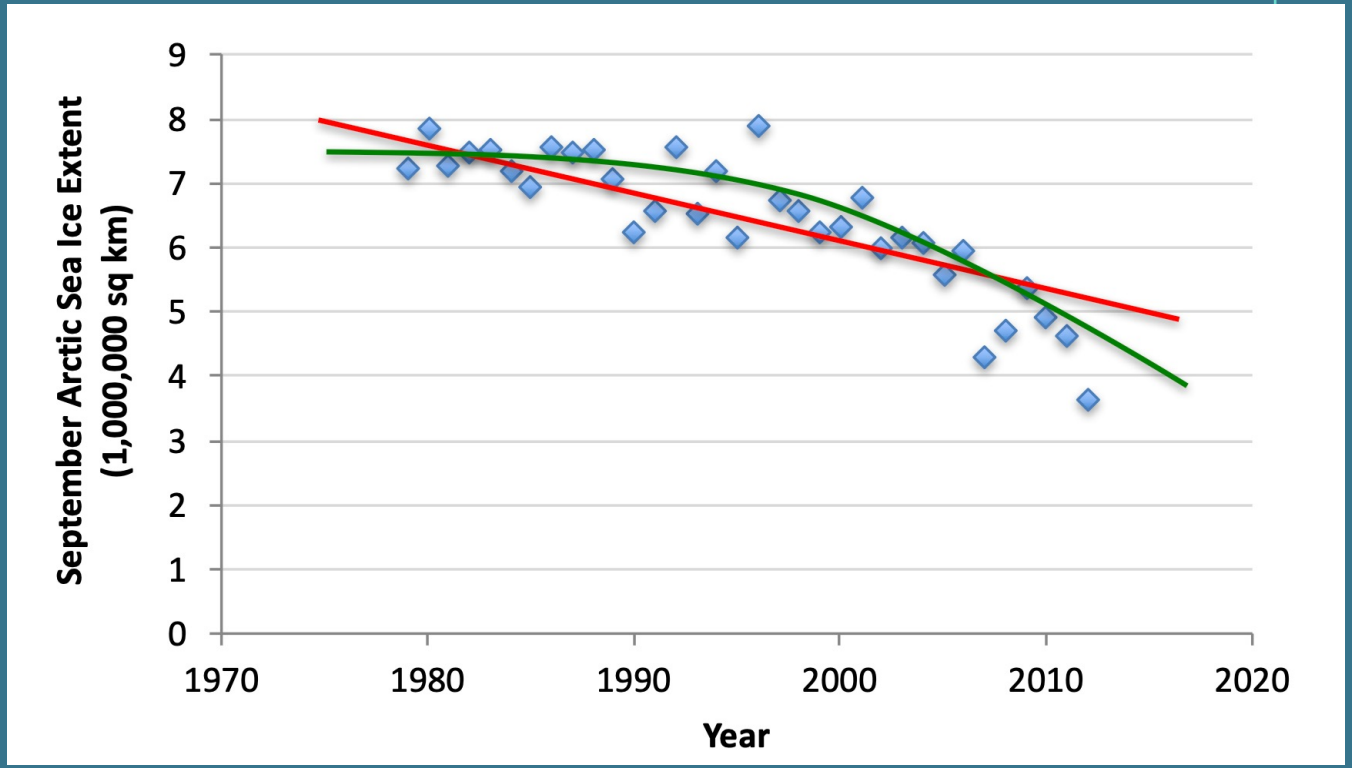
Reinforcement learning

Rewards from sequence of actions

Supervised Learning: Regression

Given
$(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$
Learn a function $f(x)$ to predict y given x

- y is real-valued == regression



Supervised Learning: Regression

Stock Price Prediction

Sales Forecasting

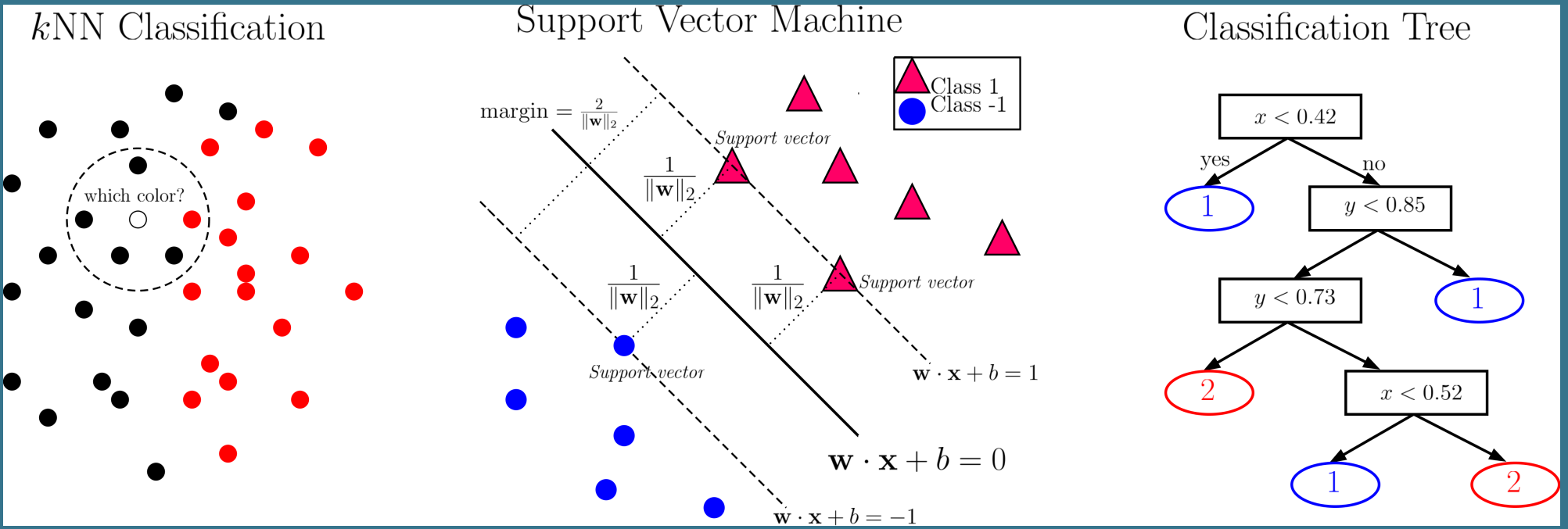
Housing prices

Risk Analysis

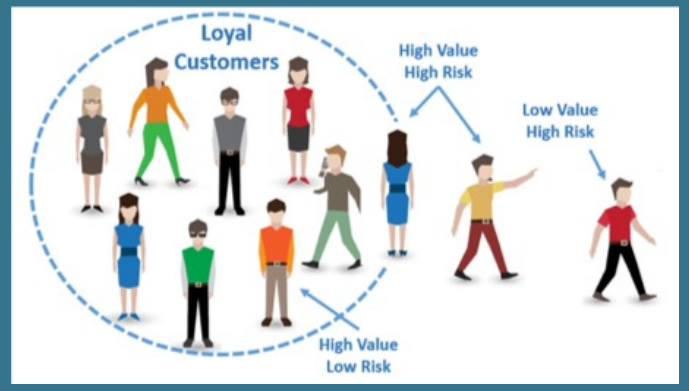
Supervised Learning: Classification

Given $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$
Learn a function $f(x)$ to predict y given x

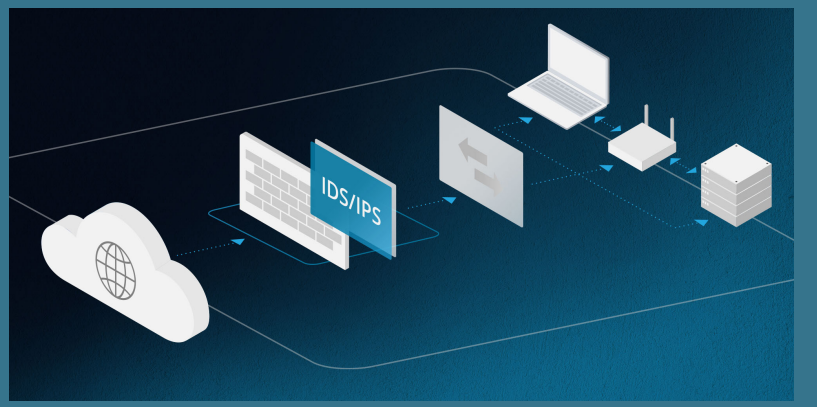
- y is categorical == classification



Supervised Learning: Classification



Customer Churn Prediction



Intrusion Detection System

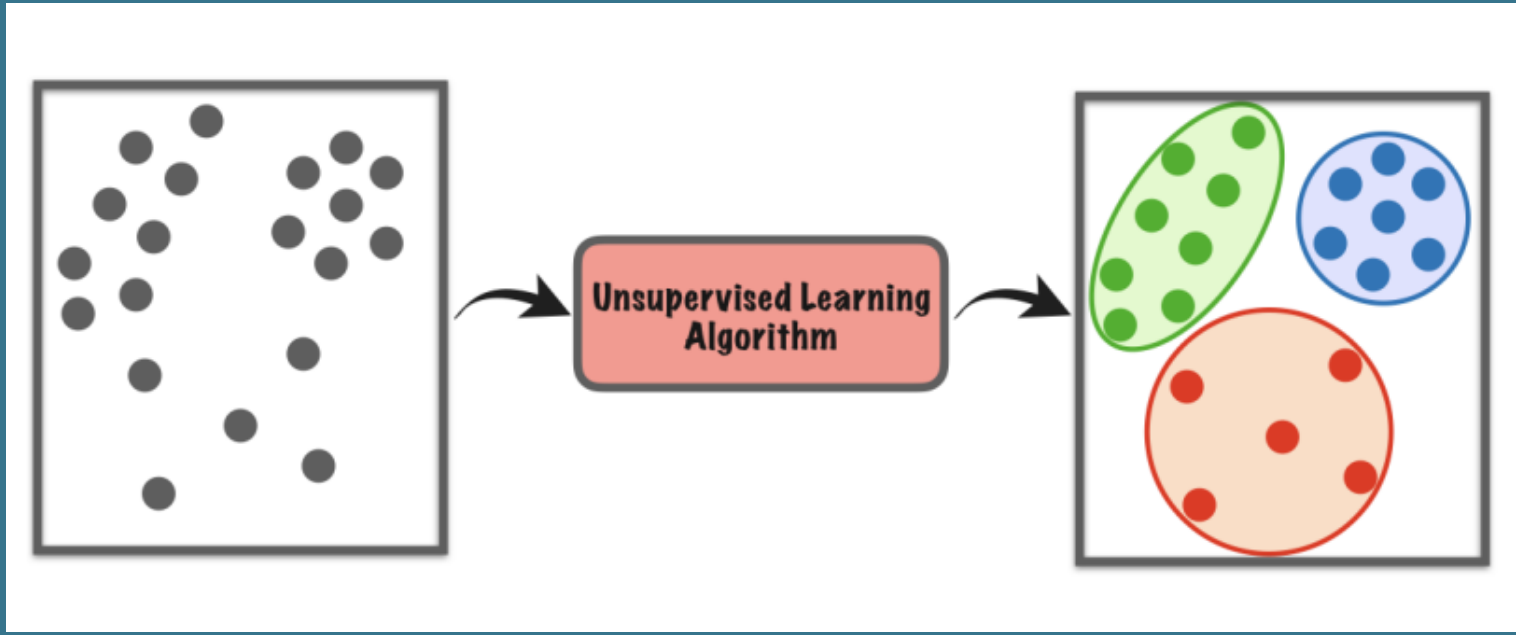
Email Spam Detection

Sentiment Analysis

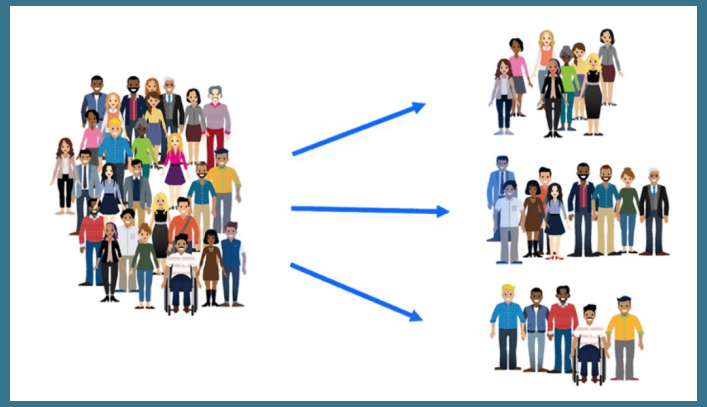
Unsupervised Learning

Given $x_1, x_2, \ldots, x_n$ (without labels)
Output hidden structure behind the x's

- E.g., clustering



Unsupervised Learning



Market segmentation

Organize computing clusters

Social network analysis

Group individuals by genetic similarity

Reinforcement Learning

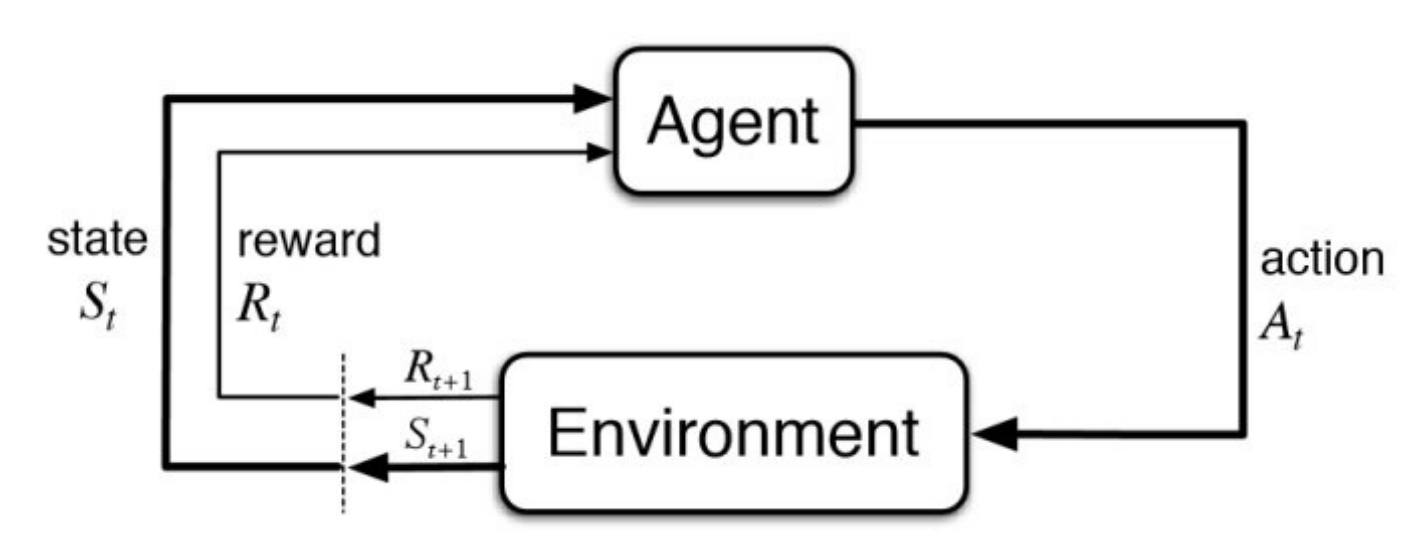
Given a sequence of states and actions with (delayed) rewards, output a policy

- Policy is a mapping from states to actions that tells you what to do in a given state

Examples:

- Trading and finance
- News recommandation
- Natural Language Processing
- Helthcare
- Gaming
- Marketing and advertising

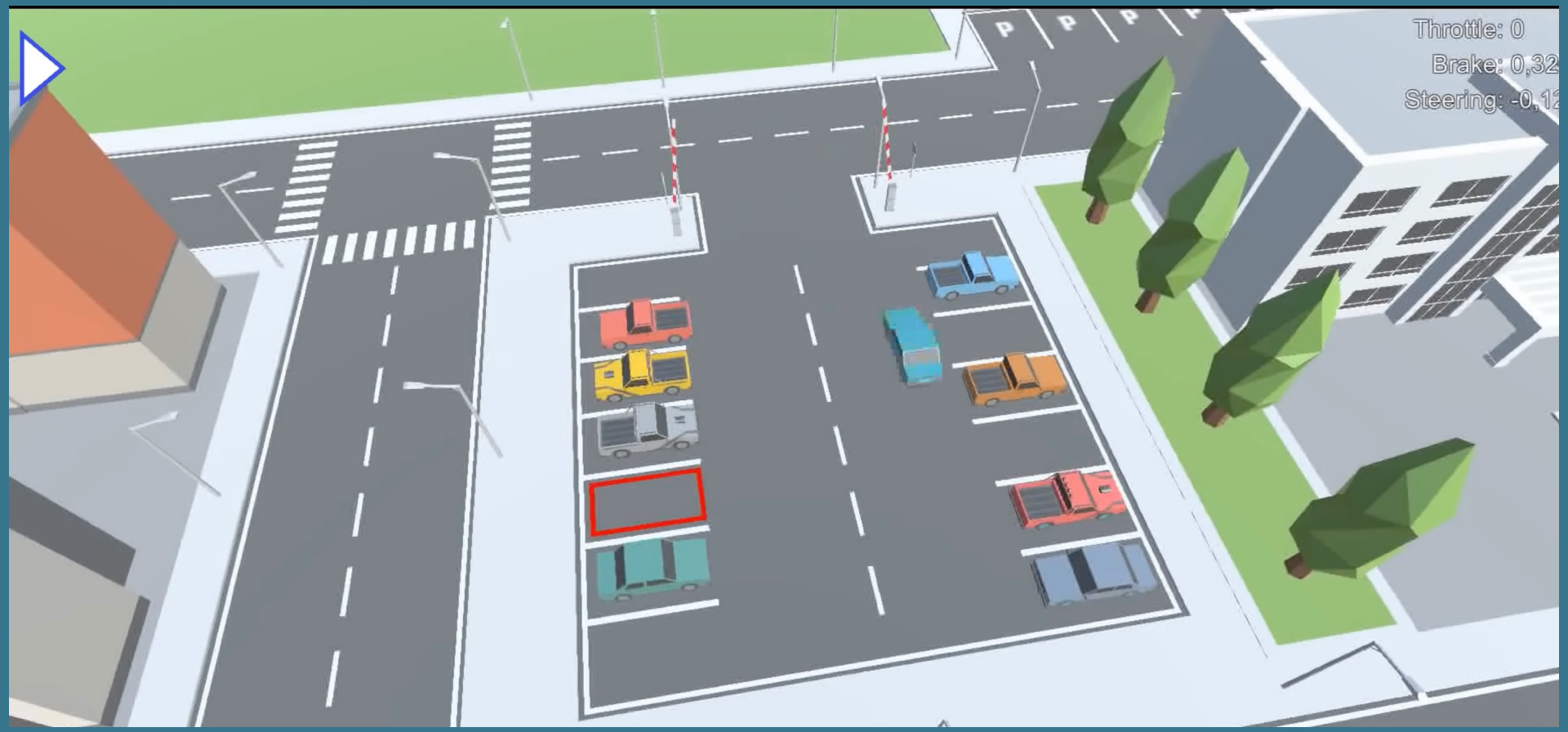
Agent-Environment Interface



Agent and environment interact at discret time steps: $t = 0,1, \dots, k$

Agent observes state at step t: $s_t \in S$
produces action at step t: $a_t \in A(S)$
gets resulting reward: $r_{t+1} \in R$
and resulting next state: s_{t+1}

AI learns to Park

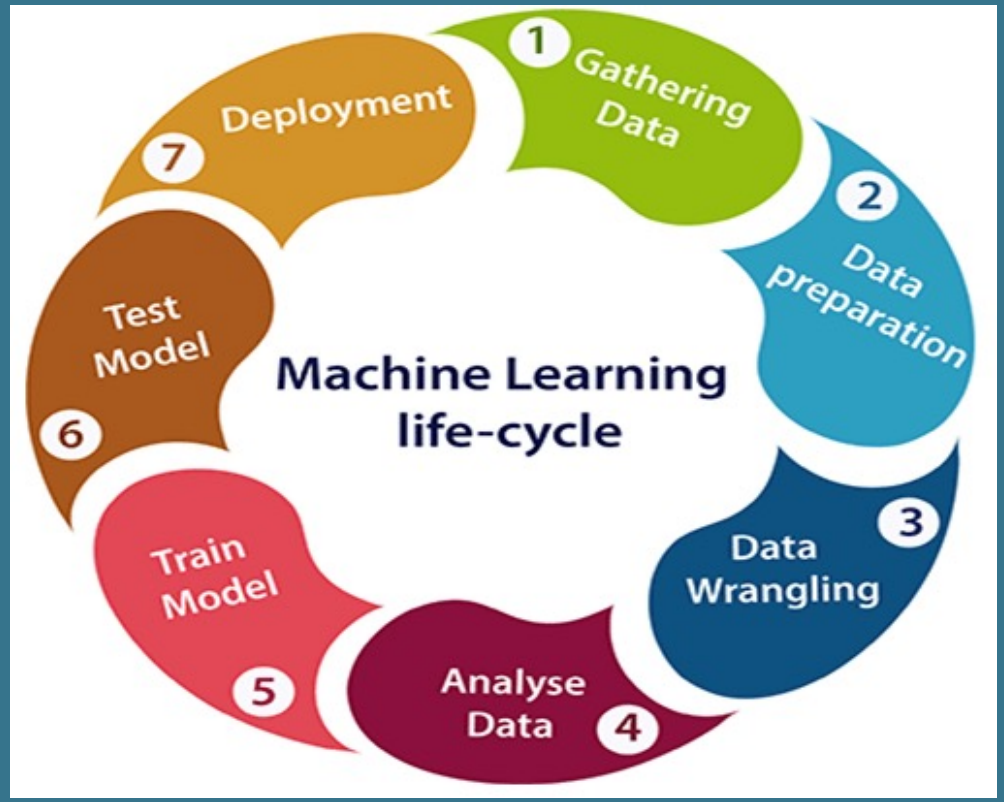


https://www.youtube.com/watch?v=VMp6pq6_QjI

ML Life Cycle

1. Gathering Data

- Identify various data sources
- Collect data
- Integrate the data



2. Data Preparation

- Data exploration
- Data pre-processing

3. Data Wrangling

- Missing values
- Duplicate data
- Invalid data
- Noise

4. Data Analysis

- Selection of analytical techniques
- Building models
- Review the result

5. Train Model

- Understand the various patterns, rules, and features

6. Test Model

- Testing the model and determines some metrics

6. Deployment

- Deploy the model in the real system

References

C. Albon, Python Machine Learning Cookbook, O'Reilly (2018)
https://cutt.ly/oJ5tSII

A.C. Müller, S. Guido, Introduction to Machine Learning with Python, O'Reilly (2017)
https://cutt.ly/6J5tB9m

A. Géron, Hands-On Machine Learning with Scikit-Learn & TensorFlow, O'Reilly (2017)
https://cutt.ly/2J5ytqJ